

日 本 国 特 許 庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

J1036 U.S. PRO  
09/821400  
03/29/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出 願 年 月 日  
Date of Application:

2000年 3月29日

出 願 番 号  
Application Number:

特願2000-091874

出 願 人  
Applicant (s):

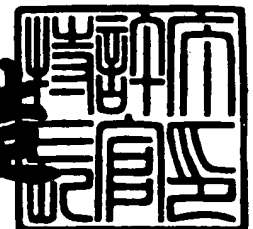
インターナショナル・ビジネス・マシーンズ・コーポレーシ  
ョン

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年12月22日

特許庁長官  
Commissioner,  
Patent Office

及 川 耕 造



【書類名】 特許願

【整理番号】 JA999315

【提出日】 平成12年 3月29日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 13/38  
G06F 15/177

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 東京基礎研究所内

【氏名】 戚 乃箴

【発明者】

【住所又は居所】 神奈川県大和市下鶴間 1 6 2 3 番地 1 4 日本アイ・ピー・エム株式会社 東京基礎研究所内

【氏名】 中村 祐一

【特許出願人】

【識別番号】 390009531

【氏名又は名称】 インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】 100086243

【弁理士】

【氏名又は名称】 坂口 博

【復代理人】

【識別番号】 100104880

【弁理士】

【氏名又は名称】 古部 次郎

【選任した代理人】

【識別番号】 100091568

【弁理士】

【氏名又は名称】 市位 嘉宏

【選任した復代理人】

【識別番号】 100100077

【弁理士】

【氏名又は名称】 大場 充

【手数料の表示】

【予納台帳番号】 081504

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9304391

【包括委任状番号】 9304392

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ウェブページ取得サービスシステム、プロバイダ、ウェブページ取得方法、記憶媒体及びプログラム伝送装置

【特許請求の範囲】

【請求項 1】 通信ネットワークに接続されたウェブページ取得サーバとユーザ端末とを備え、

前記ユーザ端末は、

種々の取得条件を指定したウェブページ取得要求を前記ウェブページ取得サーバに送信し、

前記ウェブページ取得サーバは、

前記ユーザ端末から受信した前記ウェブページ取得要求の取得条件にしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持し、

前記ウェブページソースを、前記ウェブページ取得要求の取得条件にしたがって、前記ユーザ端末に転送することを特徴とするウェブページ取得サービスシステム。

【請求項 2】 前記ユーザ端末は、

前記ウェブページ取得要求の取得条件として、前記ウェブページソースを取得するための時間的条件を設定し、

前記ウェブページ取得サーバは、

前記ウェブページ取得要求にて設定された前記時間的条件に応じて、前記ウェブページソースを取得し、前記ユーザ端末に転送することを特徴とする請求項 1 に記載のウェブページ取得サービスシステム。

【請求項 3】 前記ウェブページ取得サーバは、前記ウェブページ取得要求にて設定された前記時間的条件と前記通信ネットワークにおける通信量の多少とを考慮して前記ウェブページソースを取得するためのスケジューリングを行うことを特徴とする請求項 2 に記載のウェブページ取得サービスシステム。

【請求項 4】 前記ユーザ端末は、

前記ウェブページ取得要求の取得条件として、前記ウェブページソースを取得

する期限を指定し、

前記ウェブページ取得サーバは、

前記ウェブページ取得要求にて指定された前記取得期限の間に前記ウェブページソースを取得することを特徴とする請求項 1 に記載のウェブページ取得サービスシステム。

【請求項 5】 前記ウェブページ取得サーバは、複数の前記ユーザ端末から同一のウェブページに対する前記ウェブページ取得要求を受信した場合に、当該複数のウェブページ取得要求に対して、対応する前記ウェブページソースを 1 回取得して保持し、かつ当該ウェブページソースを、前記ウェブページ取得要求を送信した各ユーザ端末に転送することを特徴とする請求項 1 に記載のウェブページ取得サービスシステム。

【請求項 6】 インターネットへの接続サービスを提供するプロバイダにおいて、

ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける要求受付部と、

前記要求受付部にて受け付けた前記ウェブページ取得要求の取得条件にしたがって、ウェブサーバからウェブページソースを取得して保持するウェブページ取得保存部と、

前記ウェブページ取得保存部にて取得した前記ウェブページソースを、前記ウェブページ取得要求の取得条件にしたがって、当該ウェブページ取得要求を発信した前記ユーザに転送する転送制御部とを備えることを特徴とするプロバイダ。

【請求項 7】 前記転送制御部は、前記ウェブページ取得要求に従って前記ウェブページ取得保存部により取得され保持されている前記ウェブページソースを、書庫ファイルにまとめて転送することを特徴とする請求項 6 に記載のプロバイダ。

【請求項 8】 前記転送制御部は、前記ウェブページソースの転送先であるユーザの端末装置において一度に受信できるデータファイルのサイズが限られている場合に、前記ウェブページ取得保存部に保持されている前記ウェブページソースを当該端末装置が受信可能なサイズに分割した上で、各々書庫ファイルにま

とめることを特徴とする請求項 7 に記載のプロバイダ。

【請求項 9】 前記転送制御部は、前記ウェブページ取得保存部に保持されている前記ウェブページソースのリンクを、個々のウェブページソースの URL に基づく絶対的なリンクから、相対的なリンクに置き換えた上で転送することを特徴とする請求項 6 に記載のプロバイダ。

【請求項 10】 通信ネットワーク上に設けられたウェブページ取得サーバによるウェブページ取得方法において、

ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付けるステップと、

前記ウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得するためのスケジュールを作成するステップと、

作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持するステップと、

前記ウェブページソースを、前記ウェブページ取得要求の取得条件にしたがって、前記ユーザに転送するステップとを含むことを特徴とするウェブページ取得方法。

【請求項 11】 前記スケジュールを作成するステップは、

前記ウェブページ取得要求にて指定されたウェブページソースを取得するための時間的条件に従い、かつ通信ネットワークにおける通信量の多少に応じて、ウェブページ取得サーバ自身の負荷が少なくなるように、当該ウェブページ取得要求にて要求されるウェブページソースを取得する時間を決定するステップを含むことを特徴とする請求項 10 に記載のウェブページ取得方法。

【請求項 12】 前記スケジュールを作成するステップは、

前記ウェブページ取得要求を受け付けるステップにおいて、複数のユーザから同一のウェブページに対する前記ウェブページ取得要求を受け付けた場合に、各ウェブページ取得要求において指定されたウェブページを取得するための時間的条件を比較して、ウェブサーバからウェブページソースを取得する動作を行う回数ができるだけ少なくなるように前記スケジュールを作成するステップを含むことを特徴とする請求項 10 に記載のウェブページ取得方法。

【請求項13】 インターネットに接続した情報端末装置によるウェブページの取得方法において、

プロバイダに対してウェブページの取得条件を指定したウェブページ取得要求を送信するステップと、

前記ウェブページ取得要求の取得条件に含まれる時間的条件に基づいて前記プロバイダに対して前記ウェブページの転送を要求するステップと、

前記ウェブページの転送要求に応じて前記プロバイダから転送された、前記ウェブページの取得条件に合うウェブページソースを受信するステップとを含むことを特徴とするウェブページの取得方法。

【請求項14】 前記ウェブページの転送を要求するステップは、

前記プロバイダからの前記ウェブページを取得したことを知らせる通知を受信した場合に、前記ウェブページ取得要求の取得条件に含まれる前記時間的条件に関わらずに、前記プロバイダに対して前記ウェブページの転送を要求するステップを含むことを特徴とする請求項13に記載のウェブページ取得方法。

【請求項15】 前記ウェブページソースを受信するステップにおいて、

前記ウェブページを書庫ファイルの形式で受信することを特徴とする請求項13に記載のウェブページ取得方法。

【請求項16】 コンピュータに実行させるプログラムを当該コンピュータの入力手段が読取可能に記憶した記憶媒体において、

前記プログラムは、

ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける処理と、

前記ウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得するためのスケジュールを作成する処理と、

作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持する処理と、

前記ウェブページソースを、前記ウェブページ取得要求の取得条件にしたがって、前記ユーザに転送する処理とを前記コンピュータに実行させることを特徴とする記憶媒体。

【請求項 17】 コンピュータに、

ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける処理と、前記ウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得するためのスケジュールを作成する処理と、作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持する処理と、

前記ウェブページソースを、前記ウェブページ取得要求の取得条件にしたがって、前記ユーザに転送する処理とを実行させるプログラムを記憶する記憶手段と

前記記憶手段から前記プログラムを読み出して当該プログラムを送信する送信手段とを備えたことを特徴とするプログラム伝送装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、インターネットに接続してウェブサーバからウェブページを取得し閲覧する作業を支援するウェブページの取得サービスに関する。

【0002】

【従来の技術】

今日広く普及しているインターネットでは、ユーザの生活パターンに応じて、アクセスの量に時間的な偏りが生じている。すなわち、多くのインターネット・ユーザは日中に仕事を持っているため、個人的な要件でインターネットを利用するユーザは夜間にアクセスするが多い。したがって、夜間の通信量が増大し、回線が混雑することとなる。

【0003】

通信ネットワークにおいて、通信量が増大すると、これに伴ってデータ転送に要する時間が長くなる。そのため、インターネットにおいても、夜間の回線混雑時には、ユーザが所望のサイトへのリクエストを送信してから当該サイトのウェブページソースがダウンロードされるまでの待ち時間が長くなり、作業効率が悪い。



また、一般ユーザの多くは、インターネットに接続するために、電話回線を用いたダイヤルアップ接続を行う。そのため、データ転送に長時間を要することによって回線接続の時間が長くなると、それだけ回線使用料も増加することとなり、経済的にも好ましくない。

一方、プロバイダにおいても、特定の時間帯にアクセスが集中すると、サーバの負担が増大するため、サービスが低下するおそれがある。

#### 【0004】

回線が混雑している時間帯にインターネットにアクセスすることを回避する手段として、自動巡回プログラムがある。自動巡回プログラムは、ユーザのクライアントマシンに搭載され、予め設定された時刻に自動的にプロバイダにアクセスして、インターネットに接続する。そして、予め登録されたサイトに自動的にリクエストを送信して、所望のウェブページソースをダウンロードする。

この自動巡回プログラムを、回線が混雑していない時間帯に動作するように設定しておけば、ウェブページソースを取得するための待ち時間を減らすことができる。

#### 【0005】

また、プロバイダのサーバには、通常、ユーザがアクセスしたサイトのウェブページソースを一時的に保持するキャッシュ機能がある。したがって、ユーザが頻繁にアクセスするサイトのウェブページに関しては、サーバのキャッシュメモリにデータが残っている間に再度リクエストがなされたならば、サーバに保持されているウェブページソースを当該ユーザのクライアントマシンに送ることができる。この場合、プロバイダと該当サイトとの間におけるリクエストの送信及びウェブページソースの返送が行われないため、その分だけユーザの待ち時間を減らすことができる。

#### 【0006】

##### 【発明が解決しようとする課題】

上述したように、多くのユーザがインターネットにアクセスする時間帯は、そのアクセスによる通信量の増加のために、ユーザにとっては、待ち時間が長く作業効率の低下を招くと共に、回線使用料が増大し経済的でないという問題があっ

た。

また、プロバイダにおいても、アクセスの集中によりサーバの負担が増大するという問題があった。

【 0 0 0 7 】

ユーザが回線混雑時のアクセスを回避するために自動巡回プログラムを用いる場合、当該ユーザに関しては、待ち時間の増加及び回線使用料の増大を抑えることができるものの、相当数のユーザがアクセスする時間帯としてそれぞれ異なる時間を選ぶのでなければ、プロバイダにおけるサーバの負担は軽減しない。

【 0 0 0 8 】

また、上述したように、通常、プロバイダのサーバにはキャッシュ機能があるが、キャッシュメモリに蓄えられたデータは、キャッシュメモリがいっぱいになると古いものから順に機械的に削除されてしまうため、ユーザがアクセスしようとするときに、必ずしも所望のデータが残っているとは限らず、ユーザの要求に答えられるものではなかった。

【 0 0 0 9 】

そこで本発明は、回線混雑時におけるユーザのアクセスに対して、待ち時間を短縮させると共に、プロバイダにおけるサーバの負担を軽減させることを目的とする。

【 0 0 1 0 】

また、本発明は、プロバイダにおいて、ユーザの要求を考慮して効率的にサイトへのアクセスを支援するシステムを提供することを目的とする。

【 0 0 1 1 】

【課題を解決するための手段】

かかる目的のもと、本発明は、通信ネットワークに接続されたウェブページ取得サーバとユーザ端末とを備え、このユーザ端末は、種々の取得条件を指定したウェブページ取得要求をウェブページ取得サーバに送信し、このウェブページ取得サーバは、ユーザ端末から受信したウェブページ取得要求の取得条件にしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持し、このウェブページソースを、ウェブページ取得要求の取得条件にしたがっ

て、ユーザ端末に転送することを特徴としている。

【 0 0 1 2 】

ここで、このユーザ端末は、ウェブページ取得要求の取得条件として、ウェブページソースを取得するための時間的条件を設定し、このウェブページ取得サーバは、このウェブページ取得要求にて設定された時間的条件に応じて、ウェブページソースを取得し、ユーザ端末に転送することを特徴としている。

時間的条件としては、ユーザ端末からウェブページ取得サーバに対してウェブページの転送要求を行う時刻を設定することができる。

このような構成とすることにより、ユーザは、所望のウェブページを所望の時間に確実に取得できる点で好ましい。

【 0 0 1 3 】

さらにここで、このウェブページ取得サーバは、ウェブページ取得要求にて設定された時間的条件と通信ネットワークにおける通信量の多少とを考慮してウェブページソースを取得するためのスケジューリングを行うことを特徴としている。

このような構成とすることにより、回線混雑時を避けてウェブページの取得が行われるため、ウェブページ取得サーバにおける負荷を軽減することができる点で好ましい。

【 0 0 1 4 】

またここで、このユーザ端末は、ウェブページ取得要求の取得条件として、ウェブページソースを取得する期限を指定し、このウェブページ取得サーバは、このウェブページ取得要求にて指定された取得期限の間にこのウェブページソースを取得することを特徴としている。

このような構成とすることにより、所望の期間内におけるウェブページソースを取得できる点で優れている。取得期限は、その開始と終了のいずれも指定することができる。

【 0 0 1 5 】

さらにここで、ウェブページ取得サーバは、複数のユーザ端末から同一のウェブページに対するウェブページ取得要求を受信した場合に、この複数のウェブペ

ージ取得要求に対して、対応するウェブページソースを1回取得して保持し、かつこのウェブページソースを、このウェブページ取得要求を送信した各ユーザ端末に転送することを特徴としている。

このような構成とすることにより、重複したウェブページ取得要求をまとめて処理することができるため、ウェブページ取得サーバにおける負荷を軽減させることができる点で好ましい。

【0016】

また、本発明は、インターネットへの接続サービスを提供するプロバイダにおいて、ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける要求受付部と、このウェブページ取得要求の取得条件にしたがって、ウェブサーバからウェブページソースを取得して保持するウェブページ取得保存部と、このウェブページソースを、このウェブページ取得要求の取得条件にしたがって、このウェブページ取得要求を発信したユーザに転送する転送制御部とを備えることを特徴としている。

【0017】

ここで、この転送制御部は、ウェブページ取得要求に従ってウェブページ取得保存部により取得され保持されているウェブページソースを、書庫ファイルにまとめて転送することを特徴としている。

このような構成とすることにより、ユーザは取得したウェブページソースをローカルな単一のファイルとして扱うことができる点で好ましい。

【0018】

さらにここで、この転送制御部は、ウェブページソースの転送先であるユーザの端末装置において一度に受信できるデータファイルのサイズが限られている場合に、ウェブページ取得保存部に保持されているウェブページソースをこの端末装置が受信可能なサイズに分割した上で、各々書庫ファイルにまとめることを特徴としている。

このような構成とすれば、ユーザの端末装置が一度に受信できるデータファイルのサイズが小さい場合にも、ウェブページの取得サービスを受けることができる点で好ましい。

## 【 0 0 1 9 】

またここで、この転送制御部は、ウェブページ取得保存部に保持されているウェブページソースのリンクを、個々のウェブページソースのURLに基づく絶対的なリンクから、相対的なリンクに置き換えた上で転送することを特徴としている。

このような構成とすることにより、ユーザの端末装置において、ウェブページをローカルなファイルとして扱うことができる。

## 【 0 0 2 0 】

また、本発明は、通信ネットワーク上に設けられたウェブページ取得サーバによるウェブページ取得方法において、ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付けるステップと、このウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得するためのスケジュールを作成するステップと、作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持するステップと、このウェブページソースを、ウェブページ取得要求の取得条件にしたがって、ユーザに転送するステップとを含むことを特徴としている。

## 【 0 0 2 1 】

ここで、このスケジュールを作成するステップは、ウェブページ取得要求にて指定されたウェブページソースを取得するための時間的条件に従い、かつ通信ネットワークにおける通信量の多少に応じて、ウェブページ取得サーバ自身の負荷が少なくなるように、このウェブページ取得要求にて要求されるウェブページソースを取得する時間を決定するステップを含むことを特徴としている。

このような構成とすれば、回線混雑時を避けてウェブページを取得することができるため、効率的にウェブページの取得作業を行うことができる点で好ましい。

## 【 0 0 2 2 】

さらにここで、このスケジュールを作成するステップは、ウェブページ取得要求を受け付けるステップにおいて、複数のユーザから同一のウェブページに対するウェブページ取得要求を受け付けた場合に、各ウェブページ取得要求において

指定されたウェブページを取得するための時間的条件を比較して、ウェブサーバからウェブページソースを取得する動作を行う回数ができるだけ少なくなるようにスケジュールを作成するステップを含むことを特徴としている。

このような構成とすれば、重複したウェブページ取得要求をまとめて処理することができるため、効率良くウェブページの取得作業を行うことができる点で好ましい。

#### 【 0 0 2 3 】

また、本発明は、インターネットに接続した情報端末装置によるウェブページの取得方法において、プロバイダに対してウェブページの取得条件を指定したウェブページ取得要求を送信するステップと、このウェブページ取得要求の取得条件に含まれる時間的条件に基づいてプロバイダに対してウェブページの転送を要求するステップと、このウェブページの転送要求に応じてプロバイダから転送された、ウェブページの取得条件に適うウェブページソースを受信するステップとを含むことを特徴としている。

#### 【 0 0 2 4 】

ここで、このウェブページの転送を要求するステップは、プロバイダからのウェブページを取得したことを知らせる通知を受信した場合に、このウェブページ取得要求の取得条件に含まれる時間的条件に関わらずに、プロバイダに対してウェブページの転送を要求するステップを含むことを特徴としている。

このような構成とすれば、プロバイダにおいて所望のウェブページが取得された後、任意のタイミングでこのウェブページを閲覧できる点で好ましい。

#### 【 0 0 2 5 】

さらにここで、ウェブページソースを受信するステップにおいて、このウェブページを書庫ファイルの形式で受信することができる。

#### 【 0 0 2 6 】

また、本発明は、コンピュータに実行させるプログラムをこのコンピュータの入力手段が読取可能に記憶した記憶媒体において、このプログラムは、ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける処理と、このウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得

するためのスケジュールを作成する処理と、作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持する処理と、このウェブページソースを、このウェブページ取得要求の取得条件にしたがって、ユーザに転送する処理とをこのコンピュータに実行させることを特徴としている。

このような構成とすれば、このプログラムをインストールした全てのコンピュータにおいて、ウェブページ取得サービスを実施できる点で好ましい。

#### 【0027】

さらにまた、本発明は、コンピュータに、ウェブページの取得条件を指定したウェブページ取得要求をユーザから受け付ける処理と、このウェブページ取得要求の取得条件に基づいて、ウェブページソースを取得するためのスケジュールを作成する処理と、作成されたスケジュールにしたがって、通信ネットワーク上のウェブサーバからウェブページソースを取得して保持する処理と、このウェブページソースを、このウェブページ取得要求の取得条件にしたがって、ユーザに転送する処理とを実行させるプログラムを記憶する記憶手段と、この記憶手段からこのプログラムを読み出してこのプログラムを送信する送信手段とを備えたことを特徴としている。

このような構成とすれば、このプログラムをダウンロードした全てのコンピュータにおいて、ウェブページ取得サービスを実施できる点で好ましい。

#### 【0028】

##### 【発明の実施の形態】

以下、添付図面に示す実施の形態に基づいてこの発明を詳細に説明する。

図1は、本実施の形態によるウェブページ取得サービスの概念を説明する図であり、図2は、本実施の形態におけるウェブページ取得サービスを実現するシステムの全体構成を示す図である。

#### 【0029】

図1を参照すると、本実施の形態によるサービスは、ユーザ120とウェブサイト130との間に介在するプロバイダ110により提供される。図2を参照すると、本実施の形態を実現するシステムは、インターネット200に接続された

ウェブページ取得サーバ 2 1 0 と、ユーザ端末 2 2 0 と、ウェブサーバ 2 3 0 とで構成される。

図 1 において、プロバイダ 1 1 0 には図 2 に示すウェブページ取得サーバ 2 1 0 が設置される。ユーザ 1 2 0 はユーザ端末 2 2 0 を操作して、ウェブページ取得サーバ 2 1 0 により提供されるサービスを享受する。ウェブサイト 1 3 0 はウェブサーバ 2 3 0 上に構築され、種々のウェブページソースを提供する。

#### 【 0 0 3 0 】

図 1 を参照して本実施の形態によるサービスの概要を説明する。

ユーザ 1 2 0 は、プロバイダ 1 1 0 にアクセスし、当該プロバイダ 1 1 0 に対して、所望のウェブサイト 1 3 0 からウェブページを取得するための要求と当該ユーザ 1 2 0 のプロファイルとを送信する。ただし、ユーザ 1 2 0 のプロファイルはプロバイダ 1 1 0 に一度通知されていれば良く、アクセスのたびに送信する必要はない。プロバイダ 1 1 0 は、ユーザ 1 2 0 からの要求を受け付け、これに基づいてウェブサイト 1 3 0 からウェブページを取得して保持する。そして、ユーザ 1 2 0 の要求の内容に応じて、ユーザ 1 2 0 に対して通知を行い、ウェブページを転送する。

#### 【 0 0 3 1 】

図 3 は、プロバイダ 1 1 0 に設定されるウェブページ取得サーバ 2 1 0 の構成を説明する図である。

図 3 において、符号 3 1 0 は要求受付部であり、ユーザ端末 2 2 0 から送信されたウェブページ取得要求及びプロファイルを受け付けて管理する。符号 3 2 0 はウェブページ取得保存部であり、要求受付部 3 1 0 で受け付けたユーザ端末 2 2 0 からのウェブページ取得要求に応じてウェブサーバ 2 3 0 からウェブページを取得し保存する。符号 3 3 0 は転送制御部であり、ウェブページ取得保存部 3 2 0 により取得されたウェブページのユーザ端末 2 2 0 への転送を制御する。

#### 【 0 0 3 2 】

図 4 は、ユーザ 1 2 0 がユーザ端末 2 2 0 から送信するユーザプロファイルのフォーマットの例を示す図である。

ユーザプロファイルには、図示のように、ユーザ 1 2 0 に通知を行うための電



子メール (E-mail: Electronic mail) アドレス、ユーザ端末 220 が一度に受信できるデータの最大サイズ (最大転送サイズ)、転送するデータサイズが当該最大転送サイズを超える場合に分割転送をするか否かといった情報が含まれる。また、当該ユーザ 120 が定期的な取得を希望するウェブページの URL (Uniform Resource Locators)、当該ウェブページの取得頻度、取得する深さといった情報を含めることもできる。

### 【0033】

ユーザプロフィールに含まれる情報において、分割転送を行うことを指定しておけば、転送するウェブページのデータサイズが最大転送サイズを超えている場合に、当該最大転送サイズ以内のサイズにデータファイルを分割して複数回に分けて転送を行うようにプロバイダ 110 に要求できる。分割転送を行わないことを指定している場合は、最大転送サイズ分のデータのみを転送することとなる。

ウェブページの URL としては、ウェブブラウザに管理されているブックマークのデータリストを用いることもできる。

ウェブページの取得頻度とは、指定された URL のウェブページを取得する頻度を指定する。図示の例では、www.aaa.co.jp という URL を持つウェブページは月曜日、水曜日、金曜日の週 3 回取得し、www.bbb.co.jp/news という URL を持つウェブページは毎日取得すると指定されている。例えば、ニュース等を提供しているウェブサイト 130 では、記事の内容が毎日更新されるので、毎日ウェブページを取得し、あまり頻繁には更新されないウェブサイト 130 は、数日おきにウェブページを取得するように設定することができる。

取得する深さは、ウェブページのリンクをどこまで辿ってウェブページソースを取得するかを指定する。例えば、ニュースを提供するウェブサイト 130 では、ウェブページの第 1 階層に各ニュースのヘッドラインが記述され、第 2 階層に各記事の内容を記述する。そこで、このようなウェブサイト 130 でどのようなニュースがあるかだけを知るためにウェブページを取得しようとするならば、第 1 階層を取得する深さとして指定する。また、各記事の内容までを取得したいのであれば、第 2 階層までを取得する深さとして指定する。

### 【0034】

図5は、ユーザ120がユーザ端末220から送信するウェブページ取得要求のフォーマットの例を示す図である。

ウェブページ取得要求には、取得しようとするウェブページのURLのリスト、当該ウェブページをダウンロードする時刻、取得する深さといった情報が含まれる。また、当該ウェブページを取得する期限を指定することもできる。

#### 【0035】

ウェブページ取得要求に含まれる情報において、ウェブページをダウンロードする時刻を指定することによって、プロバイダ110に対していつまでに当該ウェブページをダウンロードすれば良いかを知らせることができる。

ウェブページのダウンロード時刻とは別に、当該ウェブページの取得期限を指定することにより、当該ウェブページのいつの時点における情報を取得するかを指定することができる。例えば、ニュースを提供するウェブサイト130からニュースに関するウェブページをダウンロードする場合、ウェブページの取得が早過ぎるとニュースが古くなってしまう。また、取得するウェブページの内容によっては、いつまでに取得しなければならないといった要請もあり得る。そこで、ウェブページの取得期限を指定して、適切な時期にウェブページを取得するようにすることができる。

#### 【0036】

なお、図4、図5に示したユーザプロフィール及びウェブページ取得要求に含まれる情報は例示に過ぎない。実際には、ウェブページのURLや取得する深さといった必須の情報その他、提供するサービスに応じて種々の情報を入力できるようにフォーマットを構成することができる。また、予めユーザ情報として登録しておくユーザプロフィールの情報と、必要に応じて所望の内容で送信するウェブページ取得要求の情報とを組み合わせることにより、種々のサービスを享受することができる。

#### 【0037】

要求受付部310は、上述したユーザプロフィール及びウェブページ取得要求を受け付けてユーザ120の情報を管理する。図3を参照すると、要求受付部310は、ユーザ120から受け取ったユーザプロフィールを管理するプロファイ

ル管理部 311 と、ウェブページ取得要求を管理する要求管理部 312 とを備える。プロフィール管理部 311 は、受け付けたユーザプロフィールを、ユーザ管理データベース 340 に格納して管理する。そして、ウェブページ取得要求と共に、後述するスケジューリング処理のためにウェブページ取得保存部 320 に送る。要求管理部 312 は、受け付けたウェブページ取得要求をスケジューリング処理のためにウェブページ取得保存部 320 に送る。

#### 【0038】

ウェブページ取得保存部 320 は、要求受付部 310 から受け取ったユーザプロフィール及びウェブページ取得要求に基づいてウェブページを取得するためのスケジューリング処理を行うスケジューリング部 321 と、スケジューリング部 321 により作成されたスケジュールにしたがってウェブサーバ 230 からウェブページを取得するウェブページ取得部 322 とを備える。ウェブページ取得部 322 により取得されたウェブページソースはウェブページ保存データベース 350 に格納される。

#### 【0039】

転送制御部 330 は、ユーザ 120 に対して所望のウェブページを取得したことを電子メールにより通知する通知部 331 と、ウェブページ保存データベース 350 に格納されているウェブページにおけるリンクの付け替えを行うリンク処理部 332 と、リンクの付け替えが済んだウェブページをユーザ 120 に転送する f t p / h t t p 転送部 333 とを備える。

#### 【0040】

図 6 は、本実施の形態におけるウェブページ取得サーバ 210 の動作を概略的に説明するフローチャートである。

初期条件として、ユーザプロフィールは既にプロバイダ 110 に送られており、ユーザ管理データベース 340 に格納されてプロフィール管理部 311 に管理されているものとする。そして、プロバイダ 110 によりインターネット接続サービスを受けている複数のユーザ 120 が、ユーザ端末 220 を用いてプロバイダ 110 にアクセスし、ユーザプロフィール及びウェブページ取得要求を送信したものとする。

## 【0041】

図6を参照すると、まず、要求受付部310の要求管理部312が、ウェブページ取得要求を受信し、プロファイル管理部311にて管理されているユーザプロファイルの情報を加味して要求統合処理を行う（ステップ601）。要求統合処理とは、複数のユーザ120から受け取ったウェブページ取得要求を統合して取得しようとするウェブページのリスト（以下、取得リストと称す）を作成する処理である。この際、重複している要求は一つにまとめ、重複のない取得リストを作成する。ここで、重複している要求を統合するのは、複数のユーザ120からのウェブページ取得要求に対して、当該ウェブページを1回取得することで対応することにより、ウェブページ取得サーバ210の負担を軽減するためである。

なお、ウェブページ取得要求は、複数のユーザ端末220から随時入力されるので、所定の時点でウェブページ取得要求の受付をロックして、要求統合処理を行う。

## 【0042】

図7は、二つのウェブページ取得要求を統合して重複のない取得リストを作成する例を説明する図である。

図7を参照すると、取得リスト703は、取得するウェブページごとに、当該ウェブページの取得を要求しているユーザ名、ダウンロード時刻、URL、取得する深さ、取得期限等の情報が登録されている。図示のフォーマットによれば、これらの情報が下記のように記述されている。

[ユーザ名] | ダウンロード時刻 | URL | 取得する深さ | 取得期限

## 【0043】

図7において、ユーザID「01」を持つユーザ120（以下、このユーザ120をユーザ01と表記する）のウェブページ取得要求701では、取得するウェブページのURLとしてwww.aaa.co.jpとwww.bbb.co.jpとが指定されている。また、ユーザID「02」を持つユーザ120（以下、このユーザ120をユーザ02と表記する）のウェブページ取得要求702では、取得するウェブページのURLとしてwww.aaa.co.jpとwww.ccc.comとが指定されている。すなわち、こ

これらのウェブページ取得要求 7 0 1、7 0 2 は、www.aaa.co.jp が重複している。これらを統合して取得リスト 7 0 3 を作成すると、図示のように、www.aaa.co.jp のレコードにおけるユーザ名のフィールドには、ユーザ 0 1、0 2 が記述される。ここで、www.aaa.co.jp のウェブページに関して、ユーザ 0 1 は第 1 階層のみの取得を要求するが、ユーザ 0 2 は第 2 階層までの取得を要求しているので、取得リスト 7 0 3 の深度のフィールドには、深い方の 2 階層までを取得するように記述される。また、www.aaa.co.jp に関しては、ユーザ 0 1、0 2 のいずれも取得期限を指定していないので、ダウンロード時刻までの適当な時間に当該ウェブページソースを取得するように設定することができる。

同様にして、www.bbb.co.jp のレコードにおける各フィールドには、ウェブページ取得要求 7 0 1 から取得された情報が記述され、www.ccc.com のレコードにおける各フィールドには、ウェブページ取得要求 7 0 2 から取得された情報が記述される。

#### 【 0 0 4 4 】

さらに、要求受付部 3 1 0 の要求管理部 3 1 2 は、生成された取得リスト 7 0 3 から、ウェブページの URL と当該ウェブページを要求しているユーザ 1 2 0 との対応関係を示すテーブルを作成し、転送制御部 3 3 0 に渡す。

図 8 は、図 7 の取得リスト 7 0 3 から作成された対応テーブルの例を説明する図である。

図 8 を参照すると、取得リスト 7 0 3 において同一のレコードに記述されていた URL とユーザ名とが対応付けられている。www.aaa.co.jp については、ユーザ 0 1、0 2 の両方が取得を要求しているので、URL に対応するユーザ名にユーザ 0 1、0 2 の両方が記述されている。

また、図において、URL の後にあるアスタリスク「\*」は、当該 URL のウェブページからアスタリスクの数だけリンクを辿って得られたウェブページであることを示す。すなわち、2 行目のレコードの www.aaa.co.jp/\* とあるのは、www.aaa.co.jp のウェブページから 1 階層リンクを辿って得られたウェブページとユーザとの対応を示す。

#### 【 0 0 4 5 】

図 6 において、次に、ウェブページ取得保存部 3 2 0 のスケジューリング部 3 2 1 が、取得リスト 7 0 3 に基づいてウェブページを取得するためのスケジュールを作成する（ステップ 6 0 2）。スケジュールは、取得リスト 7 0 3 に対して所定のスケジューリングルールを適用することにより作成される。スケジューリングルールとしては、システムが提供するサービス内容に応じて種々のルールを採用することができる。例えば、基本的なルールとして、

1. ダウンロード時刻の早いウェブページを優先的に取得する。
2. ネットワークにおける通信量が少ない時間帯に取得処理を行う。

といったルールを設定し、さらに、ウェブサーバ 2 3 0 がダウンしていた場合は一定時間後に再度取得処理を行う、というような特殊な態様に応じたルールを組み合わせることができる。

#### 【 0 0 4 6 】

図 9 は、図 7 に示した取得リストから上記スケジューリングルール 1 及び 2 に基づいて作成されたスケジュールの例を説明する図である。

図 9 を参照すると、作成されたスケジュール 9 0 1 は、取得するウェブページごとに、ウェブページの URL、ウェブサーバ 2 3 0 にアクセスする時刻、取得する深さ、取得期限などの情報が記述される。図示のフォーマットによれば、これらの情報が下記のように記述されている。

URL | アクセス時刻 | 取得する深さ | 取得期限

#### 【 0 0 4 7 】

図 9 のスケジュール 9 0 1 によれば、www.aaa.co.jp のウェブページソースを 1 0 時に、www.ccc.com のウェブページソースを 1 1 時に、www.bbb.co.jp のウェブページソースを 1 4 時にそれぞれ取得するようにスケジューリングされている。なお、www.ccc.com 及び www.bbb.co.jp に関しては、それぞれ取得すべき期限が指定されており、www.ccc.com は 3 月 7 日までに、www.bbb.co.jp は 3 月 5 日以後で 3 月 7 日までに取得する。ここで、取得リスト 7 0 3 において、www.bbb.co.jp のダウンロード時刻が www.ccc.com のダウンロード時刻よりも早いにも関わらず、www.ccc.com のウェブページソースを先に取得するようにスケジューリングされているのは、www.ccc.com が米国のウェブサイト 1 3 0 に存在するために、

米国のネットワーク回線が空いている時間帯に合わせて当該ウェブページソースを取得するためである。

## 【0048】

図6において、次に、ウェブページ取得保存部320のウェブページ取得部322が、スケジューリング部321により作成されたスケジュール901にしたがって、ウェブサーバ230から目的のウェブページを取得する（ステップ603）。この際、2階層以上のウェブページを取得する場合は、リンクを1階層辿るごとに当該リンク先のウェブページのURLを新たに取得することとなる。そこで、新たに得られた当該URLを加えて再帰的にスケジュール901の作成を行うことが必要である。すなわち、スケジュール901はウェブページの取得に伴ってダイナミックに更新される。

## 【0049】

図10は、ウェブページを取得しながら再帰的にスケジュール901を作成する処理を説明するフローチャートである。

図10を参照すると、まず、スケジューリング部321が、要求受付部310にて作成された取得リスト703に基づいてウェブページを取得するためのスケジュール901を作成する（ステップ1001）。そして、ウェブページ取得部322が、スケジュール901中に未処理のURLがあるかどうかを調べ、あるならば、当該URLのウェブページをスケジュール901にしたがって取得する（ステップ1002、1003）。

次に、スケジュール901の当該ウェブページにおける取得する深さを参照して、取得したウェブページ中のリンク先のソースを取得する必要があるかどうかを調べる（ステップ1004）。取得する必要がなければステップ1002に戻り、取得する必要がある場合は、当該リンク先のURLを取得リスト703に追加し（ステップ1005）、改めてスケジューリングを行ってステップ1002に戻る（ステップ1006）。

ステップ1002において、スケジュール901中のURLを全て取得したならば、処理を終了する。

## 【0050】

以上のようにして取得されたウェブページは、ウェブページ保存データベース350に順次格納される。ウェブページ保存データベース350は、仮想的な木構造を構築して取得したウェブページを保存する。この仮想的な木構造により、ウェブサーバ230のディレクトリ構造を再現する。

図11は、ウェブページ保存データベース350の仮想的な木構造を説明する図である。同図において、ルートノード1101にドメイン名が記述され、子ノード1102～1108にそれぞれウェブサーバ230におけるディレクトリの名前が記述される。また、各ウェブページのHTMLファイルや画像ファイルは、ウェブサーバ230におけるファイル構造と同じ構造となり、対応するディレクトリの下に保存される。

図11の例では、ルートノード1101のwww.aaa.co.jpの下に二つのディレクトリ（ノード1102、1103）と一つのファイル（index.html）とがある。URL形式で表すと、www.aaa.co.jp/services/、www.aaa.co.jp/software/及びwww.aaa.co.jp/index.htmlである。また、ノード1102のディレクトリservicesの下に、さらに二つのディレクトリ（ノード1104、1105）と一つのファイル（index.html）とがある。URL形式で表すと、www.aaa.co.jp/services/e-business/、www.aaa.co.jp/services/it-consl/及びwww.aaa.co.jp/services/index.htmlである。

#### 【0051】

なお、ファイルシステムの規則によって、URLに用いられるwww.aaa.co.jpのようなドメイン名はファイル名として許されないため、各ドメイン名にはユニークなIDを付ける。ユーザ120に対してウェブページソースを転送する際には、当該IDを使用することができる。また、ドメイン名に付加されたIDと対応するドメイン名の組を対応テーブルに登録する。さらに、各ウェブページにある画像ファイル名は、重複する可能性が高いため、同様にユニークなIDを付して、画像フォーマットをディレクトリ名としたディレクトリの下に置く。このIDと画像ファイルの名前の組みも対応テーブルに登録する。

図12に、図11の木構造に対して作成された対応テーブルの例を示す。

#### 【0052】



図 6 において、所定のウェブページ取得要求に対応するウェブページが取得されると、次に、転送制御部 3 3 0 の通知部 3 3 1 が、当該ウェブページ取得要求を送信したユーザ 1 2 0 に対し、ウェブページを取得したことを示す通知を電子メールにて行う（ステップ 6 0 4）。

ユーザ 1 2 0 は、ダウンロード時刻、または電子メールによるウェブページ取得の通知を確認した際に、プロバイダ 1 1 0 にアクセスしてウェブページソースのダウンロードを要求する。

#### 【 0 0 5 3 】

プロバイダ 1 1 0 のウェブページ取得サーバ 2 1 0 においてユーザ 1 2 0 からのダウンロード要求を受け取ると、要求受付部 3 1 0 のプロファイル管理部 3 1 1 が、当該ユーザ 1 2 0 のプロファイルをユーザ管理データベース 3 4 0 から読み出して、当該ダウンロード要求と共に転送制御部 3 3 0 に渡す。

転送制御部 3 3 0 における動作は、ユーザプロファイルにおいて分割転送を行うように指定してある場合と指定していない場合とで異なる。以下、簡単のため、分割転送を行う指定がない場合について説明し、次に分割転送を行う指定がある場合について説明する。

#### 【 0 0 5 4 】

図 6 において、転送制御部 3 3 0 のリンク処理部 3 3 2 が、ウェブページ保存データベース 3 5 0 から該当するウェブページソースを読み出し、リンク処理を行う（ステップ 6 0 5）。リンク処理とは、ウェブページのリンクを URL に基づく絶対的（dynamic）なリンクから、ウェブページ保存データベース 3 5 0 における当該木構造中の相対的（local）なリンクに付け替える処理である。当該ウェブページソースがユーザ端末 2 2 0 に送られた後、ユーザ 1 2 0 がウェブページのリンクを辿ろうとした際に、リンクが絶対的なリンクのままであると、インターネット 2 0 0 を介してウェブサイト 1 3 0 にアクセスしてしまうので、ローカルな動作を維持するために上記のリンク処理を行う。

#### 【 0 0 5 5 】

なお、リンク処理部 3 3 2 は、要求受付部 3 1 0 の要求管理部 3 1 2 にて作成された、ウェブページの URL と当該ウェブページを要求しているユーザ 1 2 0

との対応関係を示すテーブル（図8参照）を参照し、リンク処理の対象となったウェブページソースを要求しているユーザ120が他にいないか調べる。そして、当該ウェブページソースを要求しているユーザ120が他にいないならば、ウェブページ保存データベース350から当該ウェブページソースを消去し、対応テーブルから当該ウェブページソースに関するレコードを削除する。

図13は、図8に示した対応テーブル801に対して、ユーザ01の要求するウェブページソースが全てダウンロードされた対応テーブル801の状態を示す図である。図13を参照すると、[www.aaa.co.jp](http://www.aaa.co.jp)は、ユーザ01によってダウンロードされたものの、ユーザ02にも要求されているので、レコードは削除されていない。したがって、当該ウェブページソースもウェブページ保存データベース350に残されている。そして、当該レコードのユーザIDが「02」のみに変更されている。また、ユーザ01のみが要求していた[www.bbb.co.jp/news](http://www.bbb.co.jp/news)及び[www.bbb.co.jp/news/\\*](http://www.bbb.co.jp/news/*)のウェブページソースについては、ユーザ01のダウンロードが済んだので、当該レコードが対応テーブル801から削除される。これに伴って、当該ウェブページソース自体もウェブページ保存データベース350から消去される。

#### 【0056】

なお、ダウンロード時刻が過ぎても、ユーザ120からダウンロード要求がなされない場合、当該ウェブページソースの処分については種々の方法が考えられる。例えば、ユーザ120からのダウンロード要求があるまで保持し続けても良いし、一定期間経過後に消去しても良い。さらに、ダウンロード時刻が過ぎた時点で直ちに消去するような設定とすることもできる。これらの処分の方法については、ユーザプロファイルにて指定することができる。

#### 【0057】

図14は、[www.aaa.co.jp](http://www.aaa.co.jp)のウェブページの第2階層までをダウンロードを要求した場合におけるリンク処理の様子を示す図である。

図14において、[www.aaa.co.jp/index.html](http://www.aaa.co.jp/index.html)のウェブページ及び当該ウェブページ中に記述された二つのリンク先である[www.aaa.co.jp/services/index.html](http://www.aaa.co.jp/services/index.html)及び[www.aaa.co.jp/software/index.html](http://www.aaa.co.jp/software/index.html)のウェブページがダウンロードされる

。そこで、リンク処理部は、図 1 2 に示した対応テーブルを参照し、[www.aaa.co.jp/index.html](http://www.aaa.co.jp/index.html)の中の絶対リンクを相対リンクに直す。図示の例では、“<http://www.aaa.co.jp/services/>”の記述が“[site12/services/index.html](http://www.aaa.co.jp/services/index.html)”に変更されている。

これに対し、[www.aaa.co.jp/services/index.html](http://www.aaa.co.jp/services/index.html)及び[www.aaa.co.jp/software/index.html](http://www.aaa.co.jp/software/index.html)のウェブページにあるリンク先、すなわち、[www.aaa.co.jp/services/](http://www.aaa.co.jp/services/)及び[www.aaa.co.jp/software/](http://www.aaa.co.jp/software/)のディレクトリ（図 1 1 のノード 1 1 0 2 及びノード 1 1 0 3）以下のウェブページはダウンロードされないため、絶対リンクは変更されない。図示の例では、“<http://www.aaa.co.jp/services/e-business/>”及び“<http://www.aaa.co.jp/services/it-consl/>”の記述がそのまま残されている。

#### 【 0 0 5 8 】

以上のリンク処理により、[www.aaa.co.jp](http://www.aaa.co.jp)のウェブページの第 2 階層までをダウンロードしたユーザ端末 2 2 0 において、[www.aaa.co.jp](http://www.aaa.co.jp) に相当する site12 のウェブページをブラウザにて表示し、リンク先を呼び出すと、当該ウェブページと共にダウンロードされたウェブページが表示される。すなわち、ローカルな動作となる。そしてさらに、第 2 階層より先のリンク先を呼び出そうとする場合に、初めてインターネット 2 0 0 に接続し該当するウェブサイト 1 3 0 にアクセスすることとなる。

#### 【 0 0 5 9 】

図 6 において、上記リンク処理が行われた後、転送制御部 3 3 0 の [f t p / h t t p](http://www.aaa.co.jp) 転送部 3 3 3 が、リンク処理の施されたウェブページソースをアーカイブして（書庫ファイルに変換して）ユーザ 1 2 0 へ転送する（ステップ 6 0 6）。

図 1 5 は、ユーザ 1 2 0 へ転送されるウェブページソースの木構造を示す図である。

図 1 5 には、ユーザ ID 「0 1」を持つユーザ 1 2 0 が、[www.aaa.co.jp](http://www.aaa.co.jp)に相当する site12 を含むいくつかのウェブページソースをダウンロードする場合の木構造である。site12 のウェブページソースに関しては、第 3 階層までダウンロードする。図示のように、ドメイン名をユニークな ID に置き換えた他は、ウェブ

ページ保存データベース350に保存した時に使うディレクトリ構造をそのまま使用している。ユーザ120に転送されるウェブページソースは、図示の木構造におけるディレクトリの下にある全てのファイルである。

#### 【0060】

ユーザプロファイルにおいて、分割転送を行うことを指定されている場合は、ユーザ120からウェブページソースのダウンロード要求を受け取った後の動作において、分割転送の要否を確認し、必要ならば分割転送を行う。

図16は、分割転送の指定がなされている場合の転送制御部330の動作を説明するフローチャートである。

図16を参照すると、転送制御部330は、要求受付部310からユーザ120のダウンロード要求と当該ユーザ120のプロファイルとを受け取ると（ステップ1601）、まず、転送するウェブページソースのデータサイズが最大転送サイズを越えているかどうかを判断する（ステップ1602）。ウェブページソースのデータサイズが最大転送サイズを超えていなければ、図6に示した動作と同様に、リンク処理部332が当該ウェブページソースのリンクを付け替え、ftp/http転送部333が当該ウェブページソースをユーザ120に転送する（ステップ1603、1604）。

#### 【0061】

一方、ウェブページソースのデータサイズが最大転送サイズを超えているならば、まず、ウェブページソースを当該ユーザ端末220の最大転送サイズごとに分割する。この際、リンクが途切れないようにするため、できるだけ深さ方向の接続を維持するように分割する。そして、分割した各データファイルのリスト（分割ファイルリスト）を作成する（ステップ1605）。

この分割処理の対象は、図15に示したようなウェブページソースの木構造であるため、分割された各データファイルは、元の木構造の部分木となっている。

#### 【0062】

次に、分割ファイルリスト中に未処理のデータファイルがあるかどうかを調べ（ステップ1606）、あるならば、当該データファイルの一つに対してリンク処理部332がリンク処理を行い（ステップ1607）、当該データファイルを

アーカイブしてユーザ 1 2 0 へ転送する（ステップ 1 6 0 8）。

分割ファイルリスト中の未処理のデータファイルに対して順次ステップ 1 6 0 7、1 6 0 8 の処理を行い、分割された全てのデータファイルを転送し終えたならば、処理を終了する（ステップ 1 6 0 6）。

#### 【 0 0 6 3 】

分割して転送されたファイルは、複数のデータファイルのままであっても良いし、ユーザ端末 2 2 0 に受信された後に単一のファイルにまとめられるようにしても良い。単一のファイルにまとめることとすれば、ユーザ端末 2 2 0 においてローカルなファイルとして扱う場合にも、リンクが途切れてしまうことがない。

#### 【 0 0 6 4 】

以上のようにして、ユーザ 1 2 0 は、所望のウェブページを集めたデータファイルをまとめてプロバイダ 1 1 0 からダウンロードできるため、当該ウェブページを閲覧するために各ウェブサイト 1 3 0 にアクセス要求を送信する必要がない。また、この時点では、ウェブページソースのダウンロードにおいても、ユーザ端末 2 2 0 とウェブページ取得サーバ 2 1 0 との間でのみデータ転送が行われ、ウェブサーバ 2 3 0 とウェブページ取得サーバ 2 1 0 との間のデータ転送は発生しない。したがって、ユーザ 1 2 0 において、ウェブページを閲覧するための待ち時間が大幅に短縮されることとなる。

#### 【 0 0 6 5 】

また、プロバイダ 1 1 0 においては、予めユーザ 1 2 0 のウェブページ取得要求を受け付けておき、ネットワーク回線の通信量が少ない時間帯にウェブページを取得するためのウェブサーバ 2 3 0 への接続を行い、かつ複数のユーザ 1 2 0 による共通のウェブページ取得要求に関しては、それらをまとめて 1 回のアクセスで対応することができる。したがって、プロバイダ 1 1 0 におけるサーバの負担が大幅に軽減されることとなる。

#### 【 0 0 6 6 】

##### 【発明の効果】

以上説明したように、本発明によれば、回線混雑時におけるユーザのアクセスに対して、待ち時間を短縮させると共に、プロバイダにおけるサーバの負担を軽

減させるサービスを提供することができる。

【 0 0 6 7 】

また、本発明によればプロバイダにおいて、ユーザの要求を考慮して効率的にサイトへのアクセスを支援するシステムを提供することができる。

【図面の簡単な説明】

【図 1】 本実施の形態によるウェブページ取得サービスの概念を説明する図である。

【図 2】 本実施の形態におけるウェブページ取得サービスを実現するシステムの全体構成を示す図である。

【図 3】 プロバイダに設定されるウェブページ取得サーバの構成を説明する図である。

【図 4】 ユーザがユーザ端末から送信するユーザプロファイルのフォーマットの例を示す図である。

【図 5】 ユーザがユーザ端末から送信するウェブページ取得要求のフォーマットの例を示す図である。

【図 6】 本実施の形態におけるウェブページ取得サーバの動作を概略的に説明するフローチャートである。

【図 7】 二つのウェブページ取得要求を統合して重複のない取得リストを作成する例を説明する図である。

【図 8】 図 7 の取得リストから作成された対応テーブルの例を説明する図である。

【図 9】 図 7 に示した取得リストから作成されたスケジュールの例を説明する図である。

【図 1 0】 ウェブページを取得しながら再帰的にスケジュール 9 0 1 を作成する処理を説明するフローチャートである。

【図 1 1】 ウェブページ保存データベースの仮想的な木構造を説明する図である。

【図 1 2】 図 1 1 の木構造に対して作成された対応テーブルの例を示す図である。

【図 13】 図 8 の対応テーブルにおいて、ユーザ 01 の要求するウェブページソースが全てダウンロードされた状態を示す図である。

【図 14】 [www.aaa.co.jp](http://www.aaa.co.jp) のウェブページの第 2 階層までをダウンロードを要求した場合におけるリンク処理の様子を示す図である。

【図 15】 ユーザへ転送されるウェブページソースの木構造を示す図である。

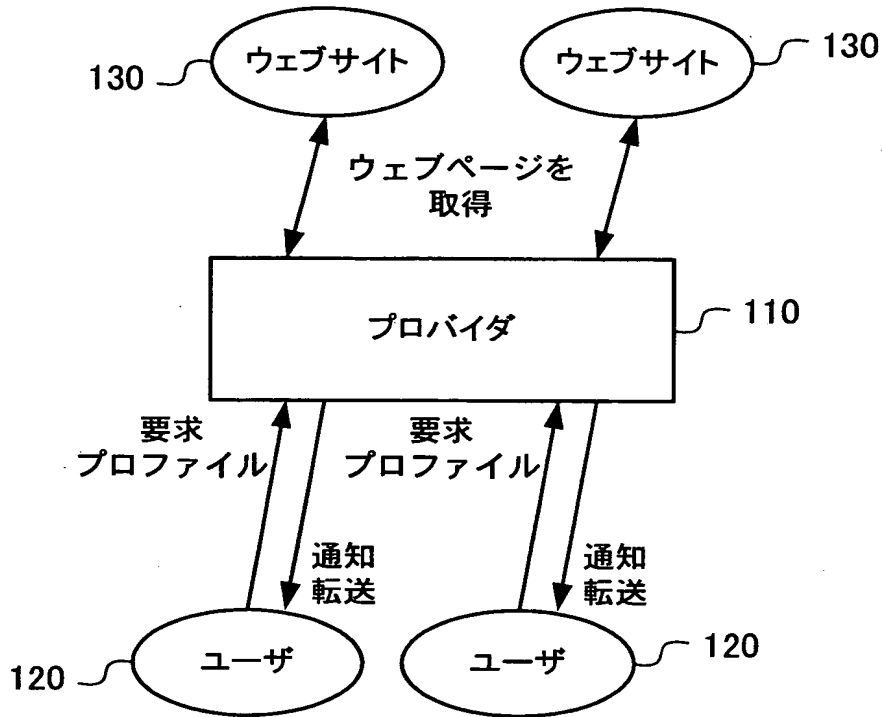
【図 16】 分割転送の指定がなされている場合の転送制御部の動作を説明するフローチャートである。

【符号の説明】

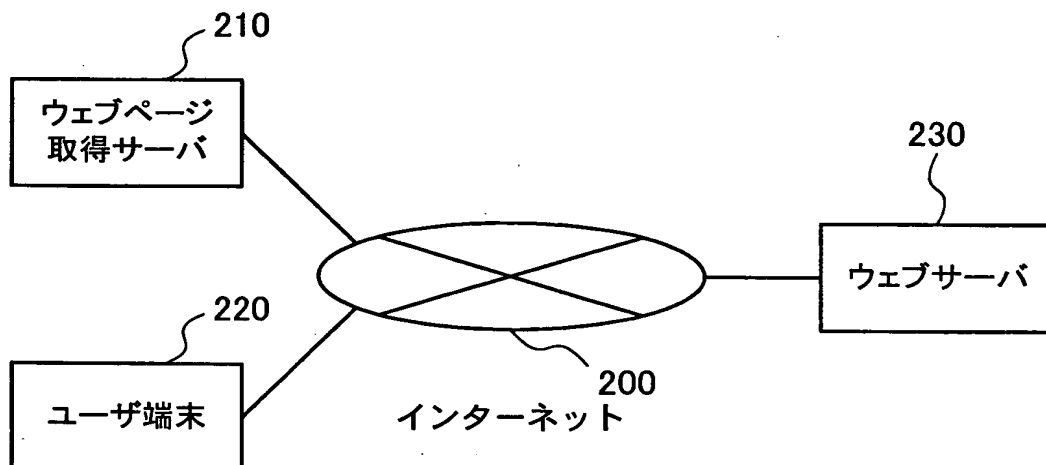
110…プロバイダ、120…ユーザ、130…ウェブサイト、200…インターネット、210…ウェブページ取得サーバ、220…ユーザ端末、230…ウェブサーバ、310…要求受付部、311…プロファイル管理部、312…要求管理部、320…ウェブページ取得保存部、321…スケジューリング部、322…ウェブページ取得部、330…転送制御部、331…通知部、332…リンク処理部、333…ftp/http 転送部、340…ユーザ管理データベース、350…ウェブページ保存データベース

【書類名】 図面

【図 1】

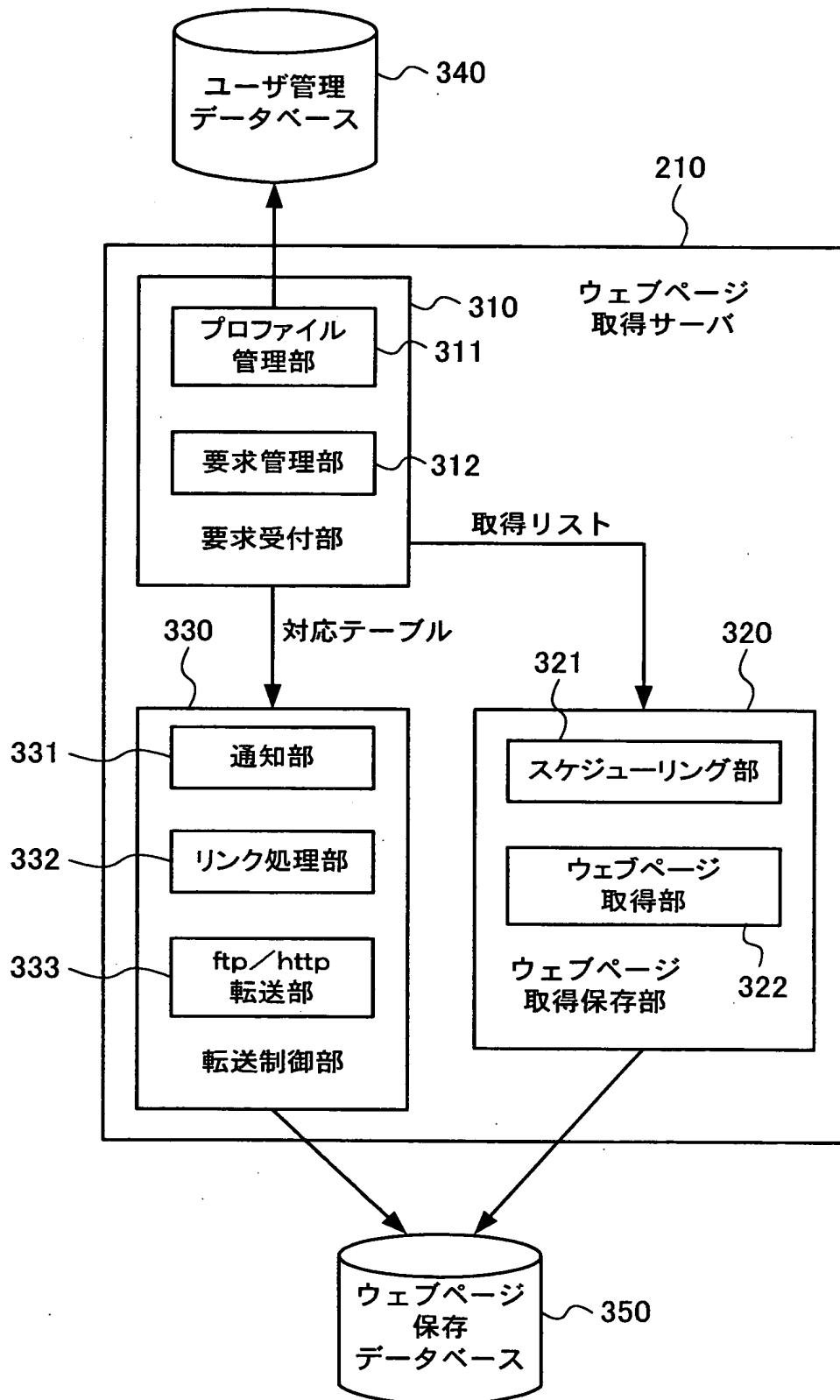


【図 2】





【図 3】



【図 4】

ユーザプロファイルのフォーマット

電子メールアドレス	abcdef@japan.aaa.com	
最大転送サイズ	200KB	
分割転送	Yes	
URL	www.aaa.co.jp www.bbb.co.jp/news	
取得頻度	www.aaa.co.jp www.bbb.co.jp/news	月曜日、水曜日、金曜日 毎日
深さ	www.aaa.co.jp www.bbb.co.jp/news	第1階層 第2階層

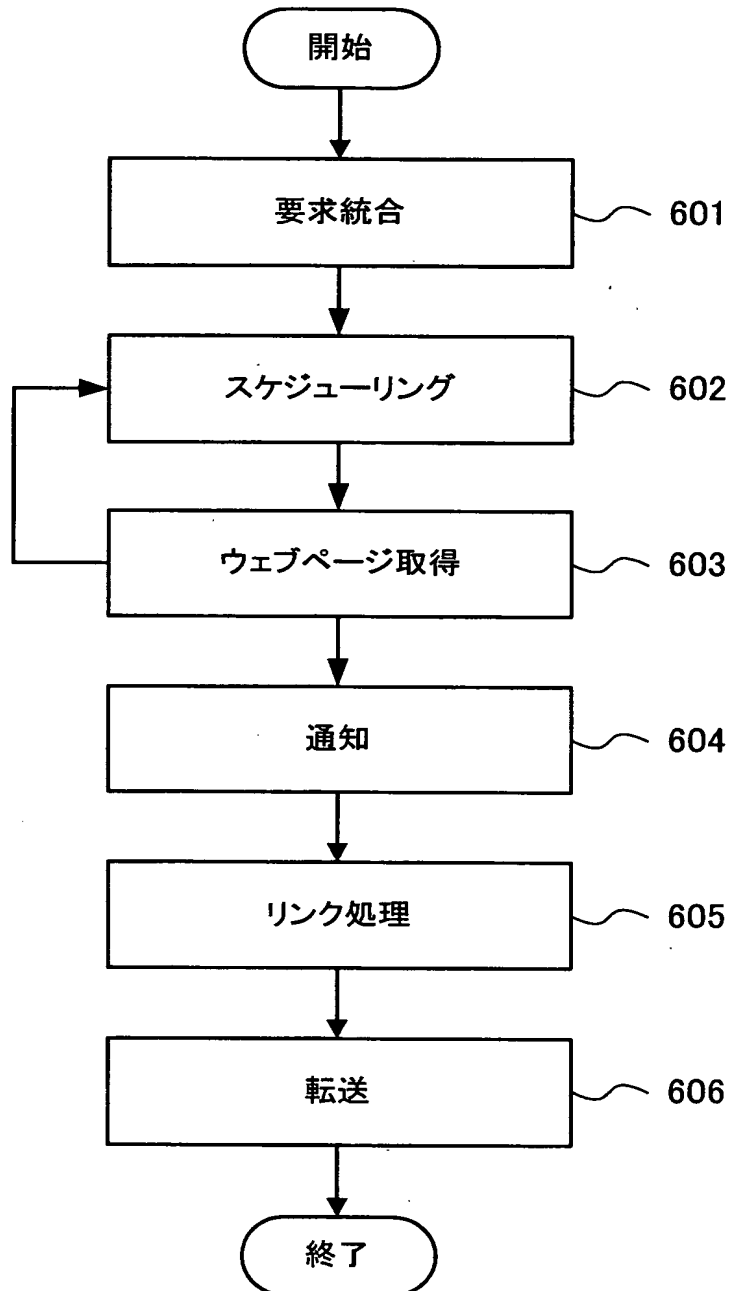
【図 5】

ウェブページ取得要求のフォーマット

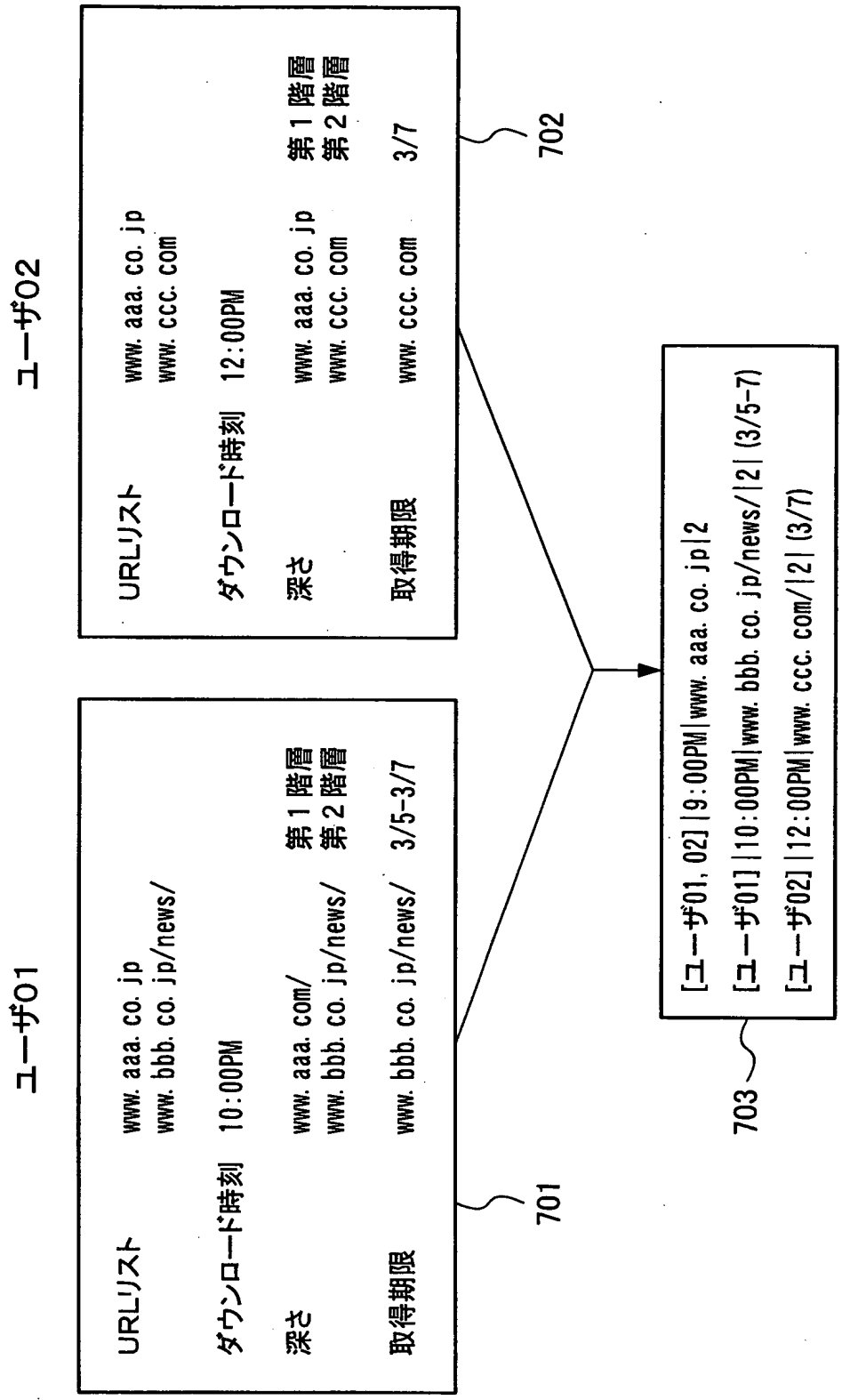
URLリスト	www.aaa.com/ www.bbb.co.jp/news/0001/11/	
ダウンロード時刻	10:00PM	
深さ	www.aaa.com/ www.bbb.co.jp/news/0001/11/	第1階層 第2階層
取得期限	3/5-3/7	

【図 6】

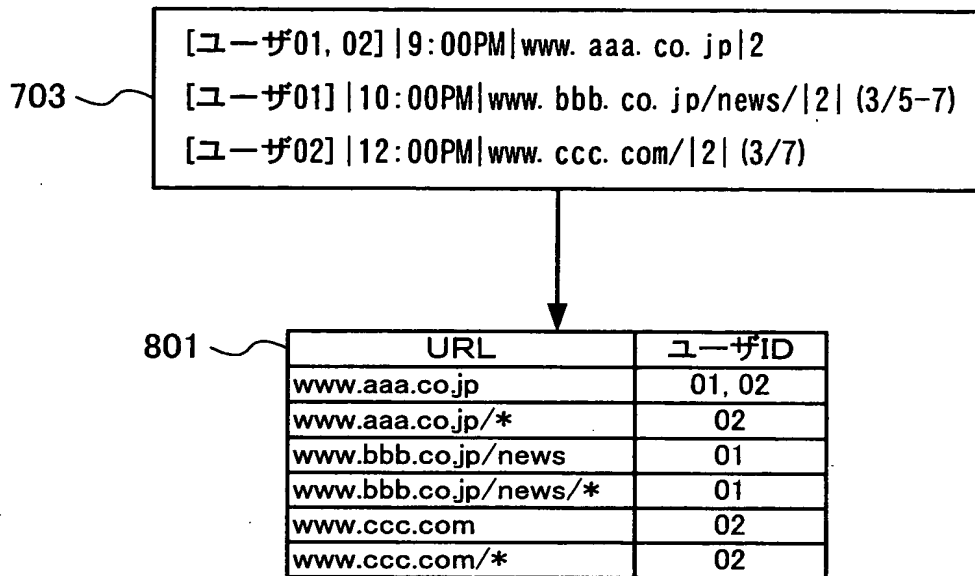
ウェブページ取得サーバの動作



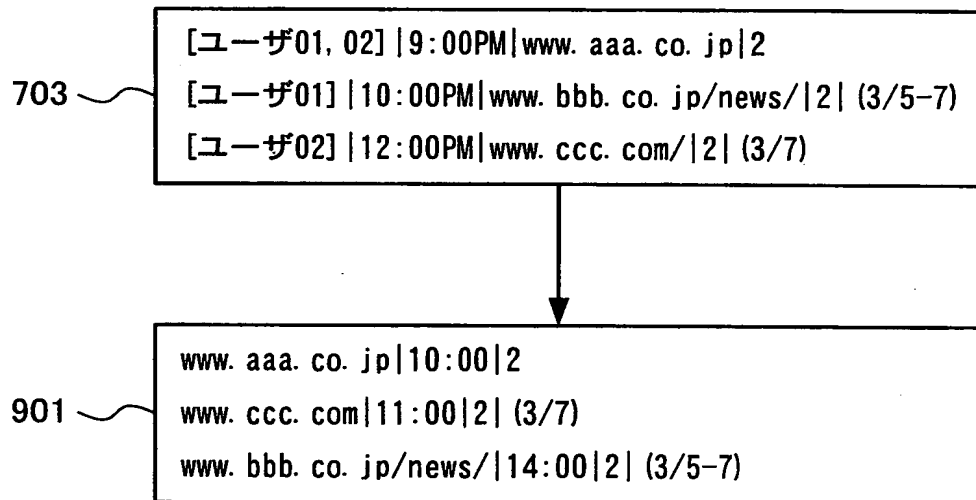
【図 7】



【図 8】

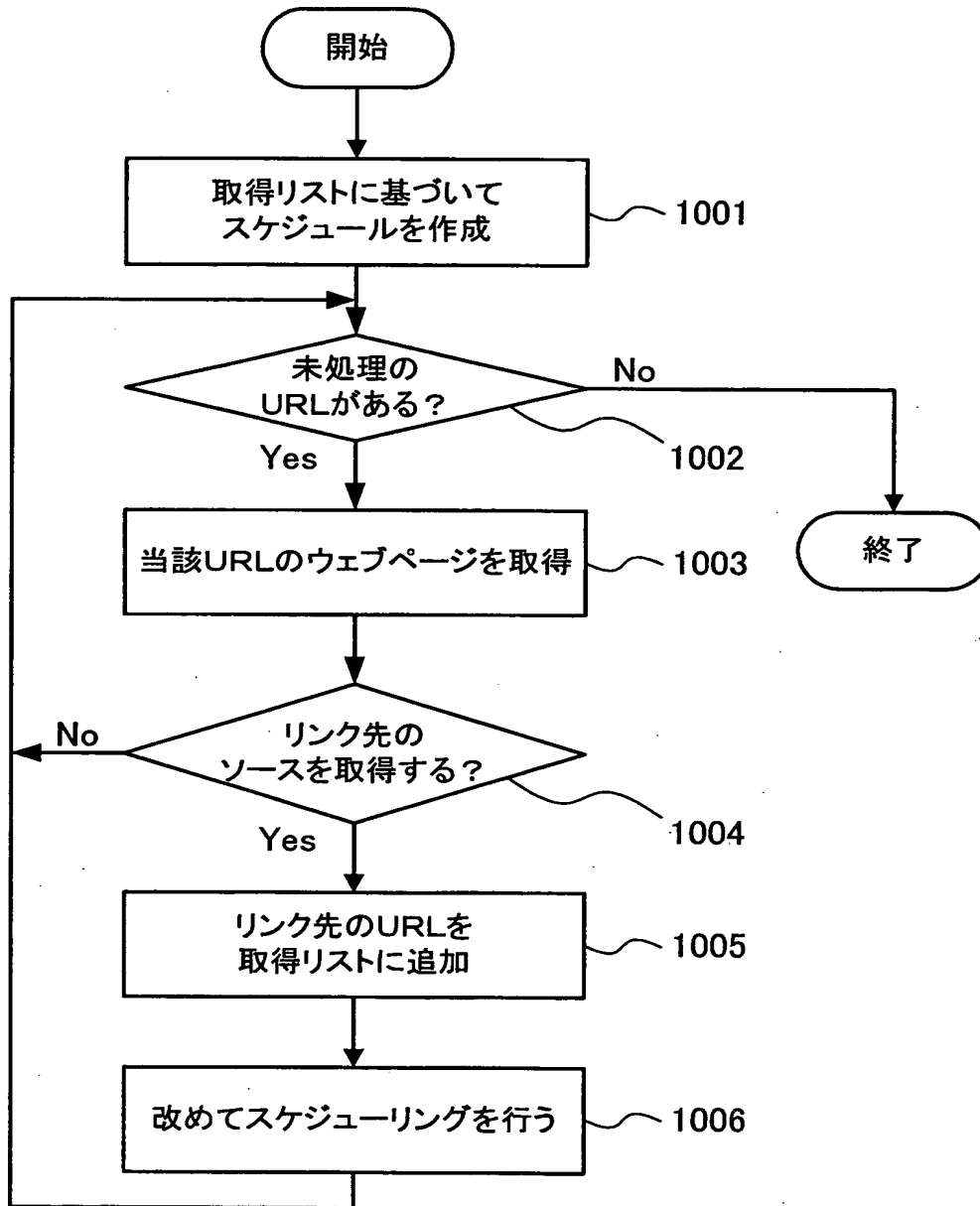


【図 9】

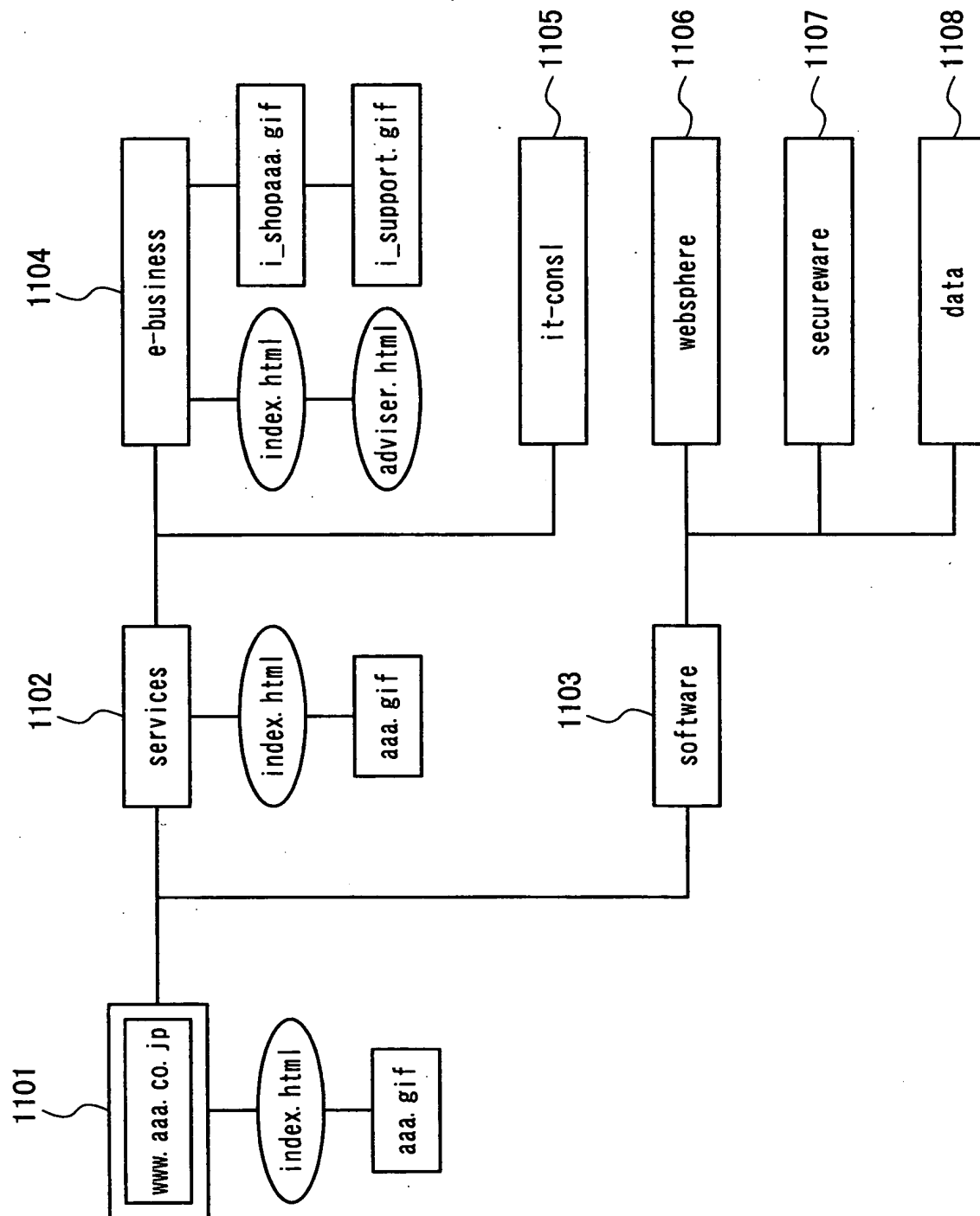


【図10】

ウェブページの取得動作



【図 11】



【図 1 2】

ドメイン名	siteID12 => www.aaa.co.jp
画像ファイル	/gif/gif03.gif => www.aaa.co.jp/gif/aaa.gif
	/gif/gif04.gif => www.aaa.co.jp/gif/i_shopaaa.gif
	/gif/gif05.gif => www.aaa.co.jp/gif/i_suport.gif

【図 1 3】

801

URL	ユーザID
www.aaa.co.jp	02
www.aaa.co.jp/*	02
www.ccc.com	02
www.ccc.com/*	02



【図 14】

http://www.aaa.co.jp/index.htmlの  
HTMLファイル

```
<HTML>
...
<A HREF="http://www.aaa.co.jp/services/">services</A>
...
</HTML>
```

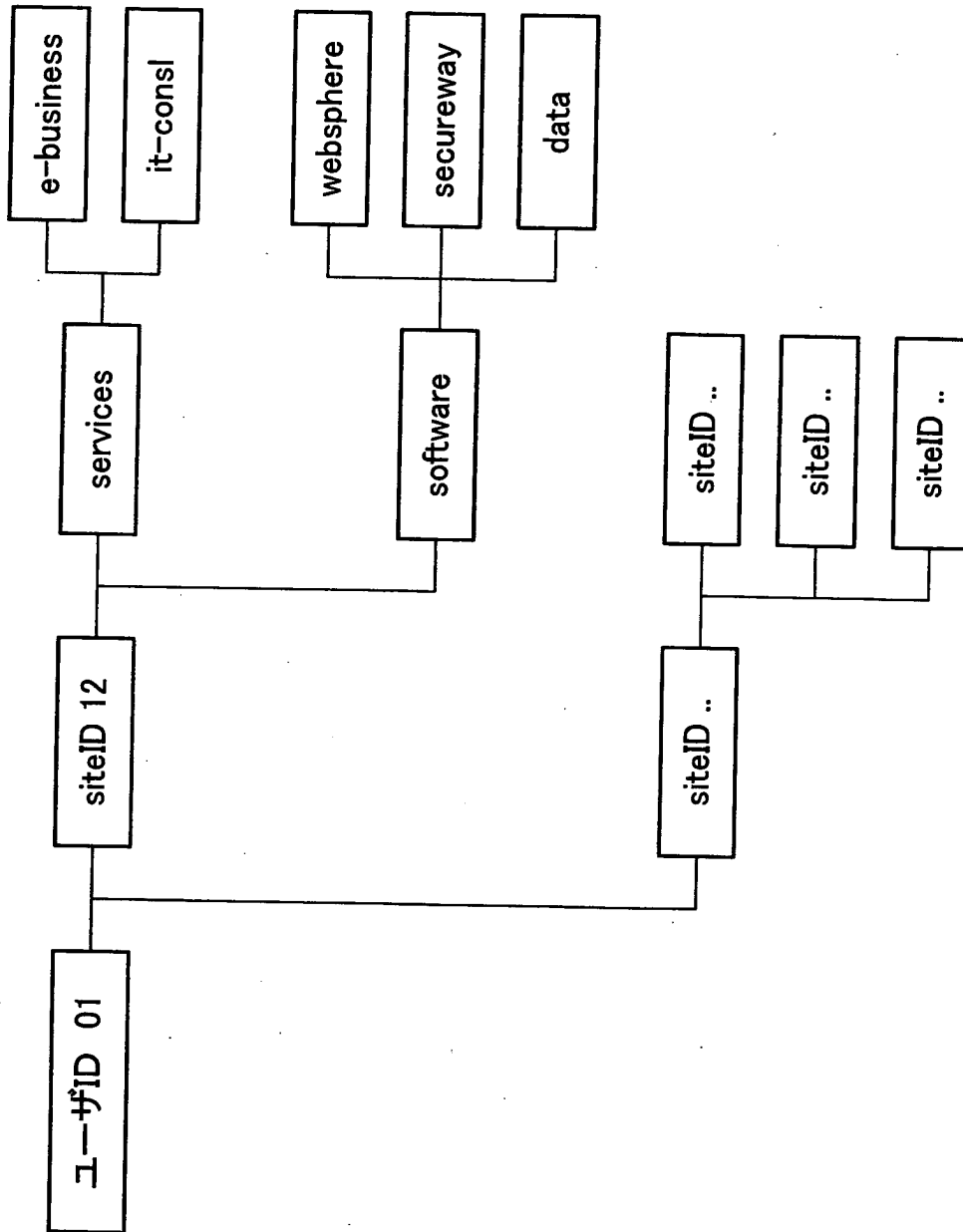


```
<HTML>
...
<A HREF="site12/services/">services</A>
...
</HTML>
```

http://www.aaa.co.jp/services/index.htmlの  
HTMLファイル

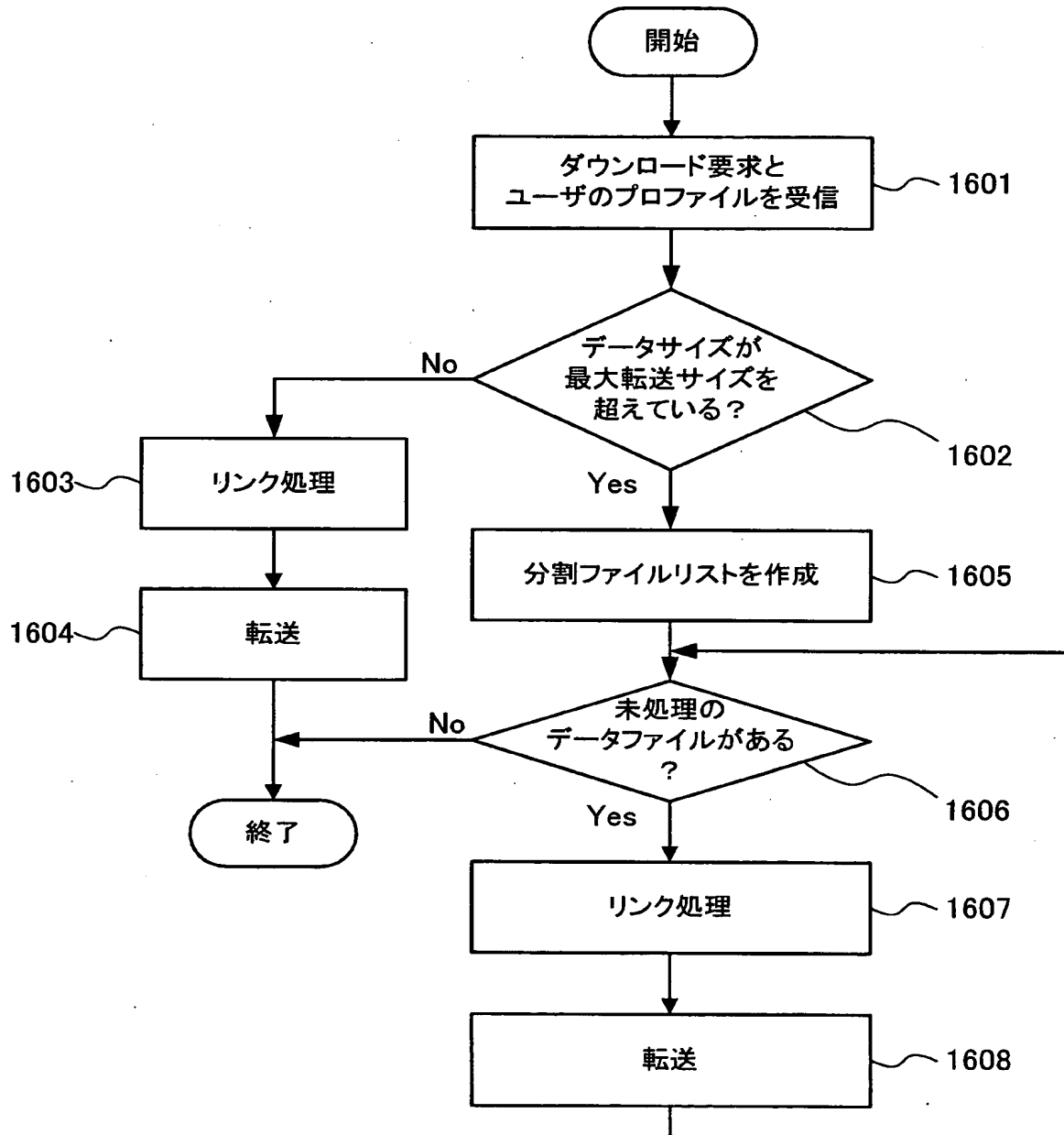
```
<HTML>
...
<A HREF="http://www.aaa.co.jp/services/e-business/">e-business</A>
<A HREF="http://www.aaa.co.jp/services/it-consl/">it-consl</A>
...
</HTML>
```

【図 15】



【図 16】

分割転送が指定されている場合の  
転送制御部の動作



【書類名】                      要約書

【要約】

【課題】    回線混雑時におけるユーザのアクセスに対して、待ち時間を短縮させると共に、プロバイダにおけるサーバの負担を軽減させる。

【解決手段】    通信ネットワークに接続されたウェブページ取得サーバ 2 1 0 とユーザ端末 2 2 0 とを備え、このユーザ端末 2 2 0 は、種々の取得条件を指定したウェブページ取得要求をウェブページ取得サーバ 2 1 0 に送信し、このウェブページ取得サーバ 2 1 0 は、ユーザ端末 2 2 0 から受信したウェブページ取得要求の取得条件にしたがって、通信ネットワーク上のウェブサーバ 2 3 0 からウェブページソースを取得して保持し、このウェブページソースを、ウェブページ取得要求の取得条件にしたがって、ユーザ端末 2 2 0 に転送する。

【選択図】                      図 2

認定・付加情報

特許出願の番号	特願 2000-091874
受付番号	50000391203
書類名	特許願
担当官	濱谷 よし子 1614
作成日	平成12年 5月17日

<認定情報・付加情報>

【特許出願人】

【識別番号】	390009531
【住所又は居所】	アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)
【氏名又は名称】	インターナショナル・ビジネス・マシーンズ・コーポレーション

【代理人】

【識別番号】	100086243
【住所又は居所】	神奈川県大和市下鶴間1623番地14 日本アイ・ビー・エム株式会社 大和事業所内
【氏名又は名称】	坂口 博

【復代理人】

【識別番号】	100104880
【住所又は居所】	東京都港区赤坂7-10-9 第4文成ビル202 セリオ国際特許事務所
【氏名又は名称】	古部 次郎

【選任した代理人】

【識別番号】	100091568
【住所又は居所】	神奈川県大和市下鶴間1623番地14 日本アイ・ビー・エム株式会社 大和事業所内
【氏名又は名称】	市位 嘉宏

【選任した復代理人】

【識別番号】	100100077
【住所又は居所】	東京都港区赤坂7-10-9 第4文成ビル202 セリオ国際特許事務所
【氏名又は名称】	大場 充

出 願 人 履 歴 情 報

識別番号 [390009531]

1. 変更年月日 1990年10月24日  
[変更理由] 新規登録  
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)  
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション
  
2. 変更年月日 2000年 5月16日  
[変更理由] 名称変更  
住 所 アメリカ合衆国10504、ニューヨーク州 アーモンク (番地なし)  
氏 名 インターナショナル・ビジネス・マシーンズ・コーポレーション